Journal of Machine Learning Research 1 (2000) 1-48 Submitted 4/00; Published 10/00

Accelerated oblique random survival forests

Byron C. Jaeger BJAEGER@WAKEHEALTH.EDU
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
Winston-Salem, NC 27157, USA

Sawyer Welden SWELDEN@WAKEHEALTH.EDU
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
Winston-Salem, NC 27157, USA

Kristin Lenoir KLENOIR@WAKEHEALTH.EDU
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
Winston-Salem, NC 27157, USA

Jaime L Speiser JSPEISER@WAKEHEALTH.EDU
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
Winston-Salem, NC 27157, USA

Matthew Segar MATTHEW.SEGAR@UTSOUTHWESTERN.EDU
Division of Cardiology, Department of Internal Medicine,
University of Texas Southwestern Medical Center, Dallas

Nicholas M. Pajewski NPAJEWSK@WAKEHEALTH.EDU
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
Winston-Salem, NC 27157, USA

Editor: TBD

Abstract

The oblique random survival forest (ORSF) is an ensemble method for supervised learning that extends the random survival forest (RSF). Trees in the ORSF are grown using linear combinations of variables to create branches in the tree, whereas in the RSF a single variable is used. ORSF ensembles often have higher prediction accuracy than RSF ensembles, but the additional computational overhead of fitting ORSF ensembles limits their scope of application. In addition, few methods have been developed for interpretation of ORSF ensembles. In this article, we introduce and evaluate methods to accelerate the ORSF (that is, reduce computational overhead) and compute the importance of individual variables in the ORSF We show that our strategy to accelerate the ORSF is up to 500 times faster than existing software for ORSFs (the `obliqueRSF` R package), and that prediction accuracy of the accelerated ORSF is equivalent or superior to that of existing ORSF methods. We estimate importance of variables for the ORSF by negating each coefficient used for the given variable in linear combinations, and then computing the reduction in out-of-bag accuracy. We show with simulation that 'negation importance' can discriminate between signal and noise variables, and it outperforms several state-of-the-art variable importance techniques in this task when there is correlation among predictors.

Keywords: Random Forests, Survival, Efficient, Variable Importance

1. Introduction

Risk prediction can reduce the burden of disease by educating patients and providers and guiding strategies to prevent and treat disease in a wide range of medical domains (Moons et al., 2012a,b). The random survival forest (RSF), a supervised learning algorithm that can engage with censored outcomes, is frequently used for risk prediction. Notable characteristics of the RSF include uniform convergence of its ensemble survival function to the true population survival function when the predictor space is discrete (Ishwaran and Kogalur, 2010). In addition, software implementing the RSF is freely available, extremely efficient, and full of tools to interpret and explain the RSF (Ishwaran and Kogalur, 2019; Wright and Ziegler, 2017; Hothorn et al., 2010). However, there remains considerable potential to improve the RSF in risk prediction tasks where training samples are not large enough to guarantee asymptotic properties or predictor spaces are non-discrete (that is, predictors are continuous).

RSFs may be axis based or oblique. The axis based RSF uses a single predictor whereas the oblique RSF uses a linear combination of predictors to create branches in trees. While axis based decision boundaries are always perpendicular to the axis of the relevant predictor, linear combinations of predictors create oblique decision boundaries that are neither parallel nor perpendicular to axes of their contributing predictors. Prior work has found the oblique RSF has higher prediction accuracy than the axis based RSF in general benchmarks (Jaeger et al., 2019) and that oblique splitting is particularly effective when predictors are continuous (Menze et al., 2011). However, existing methods to implement oblique splitting typically use fully trained models in each non-leaf node to identify linear combinations

of predictors, exponentially increasing the number of operations required for the oblique RSF versus its axis based counterpart. In addition, standard methods to estimate variable importance (VI) in the RSF are less effective in the oblique RSF, and few methods have been introduced to estimate VI specifically for the oblique RSF.

The aim of this article is to improve the computational efficiency and interpretability of the oblique RSF. In a general benchmark experiment including YYYY risk prediction tasks, we show that oblique RSFs with partially trained models have equivalent or superior prediction accuracy and are orders of magnitude more efficient than oblique RSFs with fully trained models in non-leaf nodes. We introduce a method to estimate VI for oblique RSFs and compare its ability to discriminate between signal and noise variables versus standard and state-of-the-art methods. All methods proposed in this article are available in the `aorsf` R Package.

2. Related work

2.1 Axis-based and oblique random forests

After Breiman (2001) introduced the axis-based and oblique random forest (RF), numerous methods were developed to grow oblique RFs for classification or regression tasks (Menze et al., 2011; Zhang and Suganthan, 2014; Rainforth and Wood, 2015; Zhu et al., 2015; Poona et al., 2016; Qiu et al., 2017; Tomita et al., 2020; Katuwal et al., 2020). However, oblique splitting approaches for classification or regression may not generalize to survival tasks (for example, see Zhu, 2013, Section 4.5.1), and most research involving the RSF has focused on forests with axis-based trees (Wang and Li, 2017).

Building on prior research for bagging survival trees (Hothorn et al., 2004), Hothorn et al. (2006) developed an axis-based RSF in their framework for unbiased recursive partitioning, more commonly referred to as the conditional inference forest (CIF). Zhou et al. (2016) developed a rotation forest based on the CIF and Wang and Zhou (2017) developed a method for extending the predictor space of the CIF. Ishwaran et al. (2008) developed an axis-based RSF with strict adherence to the rules for growing trees proposed in Breiman (2001). Jaeger et al. (2019) developed the oblique RSF following the bootstrapping approach described in Breiman's original RF and incorporating early stopping rules from the CIF.

2.2 Variable importance

Breiman (2001) introduced permutation VI, defined for each predictor as the difference in a RF's estimated generalization error before versus after the predictor's values are randomly permuted. Strobl et al. (2007) identified bias in permutation VI driven by variable selection bias and effects induced by bootstrap sampling, and proposed an unbiased permutation VI based on unbiased recursive partitioning (see Hothorn et al. (2006)). Menze et al. (2011) introduced an approach to estimate VI for oblique RFs that computes an analysis of

variance (ANOVA) table in non-leaf nodes to obtain p-values for each predictor contributing to the node. The ANOVA VI[1] is then defined for each predictor as the number of times a p-value associated with the predictor is ≤ 0.01 while growing a forest. Lundberg and Lee (2017) introduced a method to estimate VI using SHapley Additive exPlanation (SHAP) values, which estimate the contribution of a predictor to a model's prediction for a given observation. SHAP VI is computed for each predictor by taking the mean absolute value of SHAP values for that predictor across all observations in a given set.

3. The acceleracted oblique random survival forest

Consider the usual framework for survival analysis with training data

$$\mathcal{D}_{\text{train}} = \{(T_i, \delta_i, \boldsymbol{x}_i)\}_{i=1}^{N_{\text{train}}}.$$

Here, T_i is the event time if $\delta_i = 1$ and last point of contact if $\delta_i = 0$, and $\boldsymbol{x}_i$ is a vector of predictors values. Assuming there are no ties, let $t_1 < \ldots < t_m$ denote the m unique event times in $\mathcal{D}_{\text{train}}$.

3.1 Partial training at non-leaf nodes

We propose to identify linear combinations of predictor variables in non-leaf nodes by applying Newton Raphson scoring to the partial likelihood function of the Cox regression model:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{m} \frac{e^{\boldsymbol{x}_{j(i)}^T \boldsymbol{\beta}}}{\sum_{j \in R_i} e^{\boldsymbol{x}_j^T \boldsymbol{\beta}}}, \tag{1}$$

where R_i is the set of indices, j, with $T_j \geq t_i$ (i.e., those still at risk at time t_i), and $j(i)$ is the index of the observation for which an event occurred at time t_i. The `survival` package includes documentation that outlines how to complete this estimation procedure efficiently (see Therneau, 2022, exact.nw). Briefly, a vector of estimated regression coefficients, $\hat{\beta}$, is updated in each step of the procedure based on its first derivative, $U(\hat{\beta})$, and second derivative, $H(\hat{\beta})$:

$$\hat{\beta}^{k+1} = \hat{\beta}^k + U(\hat{\beta} = \hat{\beta}^k)\, H^{-1}(\hat{\beta} = \hat{\beta}^k)$$

It is vital to cycle through iterations until a convergence threshold is met for statistical inference, but it is only necessary to complete one iteration to identify coefficients for a linear combination of predictors. Jaeger et al. (2019) identified linear combinations using penalized regression models, which supply more flexible solutions for $\hat{\beta}$ at the cost of greater computational demand.

1. Menze et al. (2011) name their method 'oblique RF VI', but we use the name 'ANOVA VI' in this article to avoid confusing Menze's approach with other approaches to estimate VI for oblique RFs.

3.2 Negation variable importance

Negation VI is similar to permutation VI in that it measures how much a model's prediction error increases when a variable's role in the model is de-stabilized. More specifically, negation VI measures the increase in an oblique RF's prediction error after flipping the sign of all coefficients linked to a variable (that is, negating them). As the magnitude of a coefficient increases, so does the probability that negating it will change the oblique RF's predictions. Since the coefficients in each non-leaf node of an oblique RFs are adjusted for the accompanying predictors, negation VI may provide better estimation of VI in the presence of correlated variables compared to standard VI techniques.

For consistency with prior VI techniques for the RSF and for its computational efficiency, we use Harrell's concordance (C)-statistic (Harrell et al., 1982) to measure change in prediction error when computing negation VI. We also note that while the current article focuses on oblique RSFs, negation VI can be applied to any oblique RF and can be applied with any applicable error function.

4. Numeric experiments

4.1 Benchmark of prediction accuracy and computational efficiency

The aim of this numeric experiment is to evaluate and compare the accelerated oblique RSF with its predecessor (the oblique RSF from the `obliqueRSF` R package) and with other machine learning algorithms for risk prediction. Inferences drawn from this experiment include equivalence and inferiority tests based on Bayesian hierarchical models.

4.1.1 Learners

We consider four classes of learners: RFs, boosting ensembles, regression models, and neural networks (Table 1). For RF learners, the number of observations required in terminal nodes was fixed at 10, the number of randomly selected predictors was the square root of the total number of predictors rounded to the nearest integer, and the number of trees in the ensemble was 500. For boosting, regression, and neural network learners, nested cross-validation was applied to tune each model, that is, assess multiple possible values for relevant model parameters and use the values that optimize the model's estimated prediction accuracy. Specifically, tuning for boosting models included identifying the number of steps to complete. For regression models, tuning was used to identify the magnitude of penalization. For neural networks, the number and density of layers was tuned.

Learner Class	Software	Learners	Description
Random Survival Forests			

Axis based	`RandomForestSRC`
	`ranger`
	`party`
	`rotsf`
	`rsfse`
		`rsf-standard`
		`rsf-extratrees`
		`cif-standard`
		`cif-rotate`
		`cif-spacextend`
			`rsf-standard` grows survival trees following Leo Breiman's original random forest algorithm with variables and cut-points selected to maximize a log-rank statistic. `rsf-extratrees` grows survival trees with randomly selected features and cut-points. `cif-standard` uses the framework of conditional inference to grow survival trees. `cif-rotate` extends `cif-standard` by applying principal component analysis to random subsets of data prior to growing each survival tree. `cif-spacextend` derives new predictors for each tree in the ensemble, separately.

Oblique	`obliqueRSF`
	`aorsf`
		`obliqueRSF-net`
		`aorsf-net`
		`aorsf-fast`
		`aorsf-cph`
		`aorsf-extratrees`
			Oblique survival trees following Leo Breiman's random forest algorithm. Linear combinations of inputs are derived using `glmnet` in `obliqueRSF-net` and `aorsf-net`, using Newton Raphson scoring for the Cox partial likelihood function in `aorsf-fast` and `aorsf-cph`, and chosen randomly from a uniform distribution in `aorsf-extratrees`. Cut-points are selected to maximize a log-rank statistic.

Boosting ensembles			
Trees	`xgboost`	`xgboost-cox` `xgboost-aft`	`xgboost-cox` maximizes the Cox partial likelihood function, whereas `xgboost-aft` maximizes the accelerated failure time likelihood function. Nested cross validation (5 folds) is applied to tune the number of trees grown, the minimum number of observations in a leaf node was 10, the maximum depth of trees was 6, and $\sqrt{p}$ variables were considered randomly for each tree split, where p is the total number of predictors.
Regression models			
Cox Net	`glmnet`	`glmnet-cox`	The Cox proportional hazards model is fit using an elastic net penalty. Nested cross validation (5 folds) is applied to tune penalty terms.
Neural networks			
Cox Time	`survivalmodels`	`nn-cox`	A neural network based on the proportional hazards model with time-varying effects. Nested cross-validation was applied to select the number of layers (from 1 to 8), the number of nodes in each layer (from $\sqrt{p}/2$ to $\sqrt{p}$), and the number of epochs to complete (up to 500). A drop-out rate of 10% was applied during training.

Table 1: Learning algorithms assessed in numeric studies

4.1.2 Evaluation of prediction accuracy

Our primary metric for evaluating the accuracy of predicted risk is the integrated and scaled Brier score (Graf et al., 1999). Consider a testing data set:

$$\mathcal{D}_{\text{test}} = \{(T_i, \delta_i, x_i)\}_{i=1}^{N_{\text{test}}}.$$

Let $\widehat{S}(t \mid x_i)$ be the predicted probability of survival up to a given prediction horizon of $t > 0$. For observation i in $\mathcal{D}_{\text{test}}$, let $\widehat{S}(t \mid \boldsymbol{x}_i)$ be the predicted probability of survival up to a given prediction horizon of $t > 0$. Define

$$\widehat{\text{BS}}(t) = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \{\widehat{S}(t \mid \boldsymbol{x}_i)^2 \cdot I(T_i \leq t, \delta_i = 1) \cdot \widehat{G}(T_i)^{-1} + [1 - \widehat{S}(t \mid \boldsymbol{x}_i)]^2 \cdot I(T_i > t) \cdot \widehat{G}(t)^{-1}\}$$

where $\widehat{G}(t)$ is the Kaplan-Meier estimate of the censoring distribution. As $\widehat{\text{BS}}(t)$ is time dependent, integration over time provides a summary measure of performance over a range of plausible prediction horizons. The integrated $\widehat{\text{BS}}(t)$ is defined as

$$\widehat{\mathcal{BS}}(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \widehat{\text{BS}}(t) dt. \tag{2}$$

In our results, t_1 and t_2 are the 25th and 75th percentile of event times, respectively. $\widehat{\mathcal{BS}}(t_1, t_2)$, a sum of squared prediction errors, can be scaled to produce a measure of explained residual variation (that is, an R^2 statistic) by computing

$$R^2 = 1 - \frac{\widehat{\mathcal{BS}}(t_1, t_2)}{\widehat{\mathcal{BS}}_0(t_1, t_2)} \tag{3}$$

where $\widehat{\mathcal{BS}}_0(t_1, t_2)$ is the integrated Brier score when a Kaplan-Meier estimate for survival based on the training data is used as the survival prediction function $\widehat{S}(t)$. We refer to this R^2 statistic as the index of prediction accuracy (Kattan and Gerds, 2018) and we scale its values by 100 to avoid unnecessary leading zero's. For example, we present 25 if R^2 is 0.25 and present 10.2 if the difference between two R^2 is 0.102.

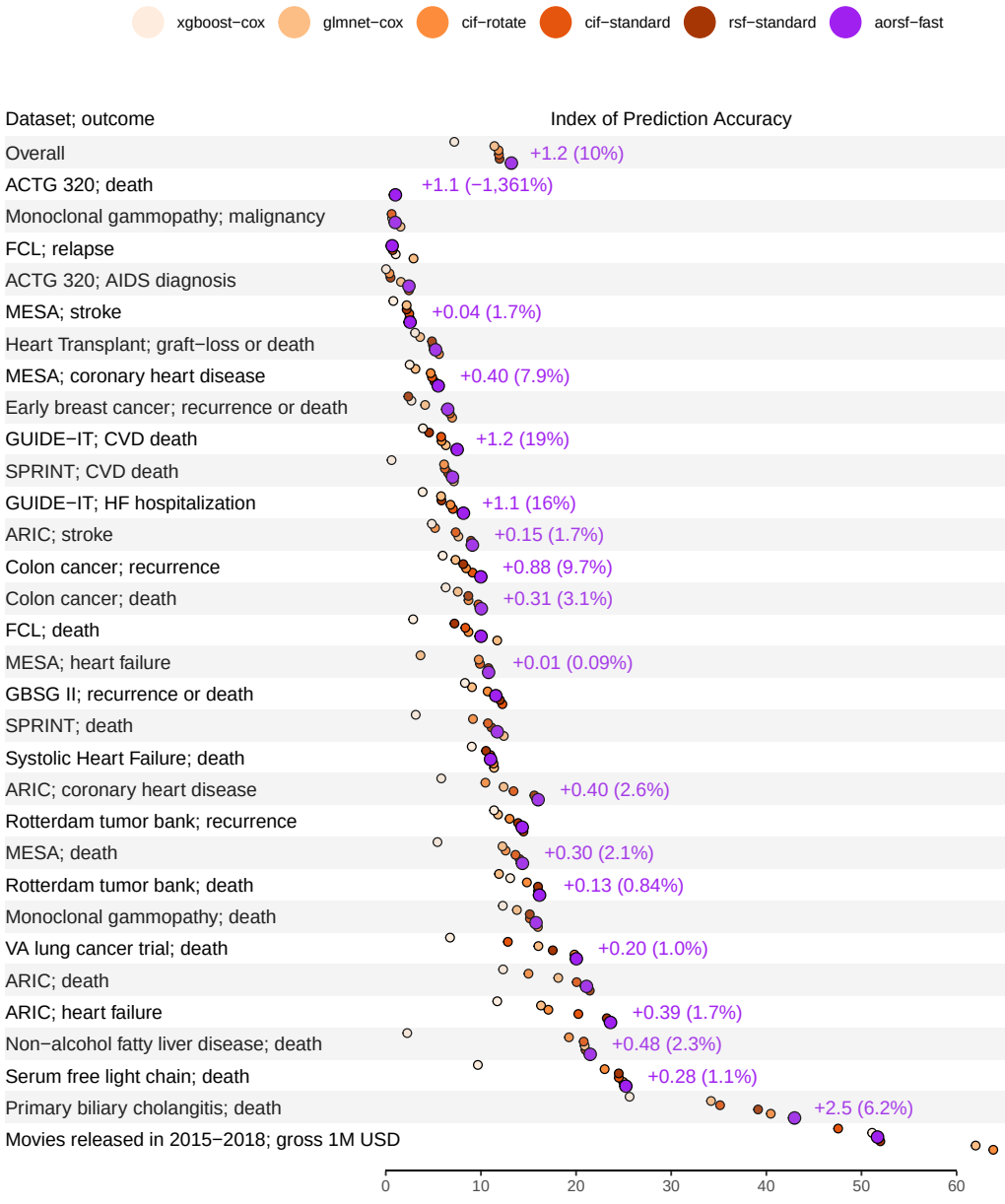
Our secondary metric for evaluating predicted risk is the time-dependent concordance (C)-statistic. Briefly, we compute the time-dependent C-statistic proposed by Blanche et al. (2013), which is interpreted as the probability that a risk prediction model will assign higher risk to a case (that is, an observation with $T \leq t$ and $\delta = 1$) versus a non-case (that is, an observation with $T > t$). Similar to the index of prediction accuracy, observations with $T \leq t$ and $\delta = 0$ only contribute to inverse proportion of censoring weights.

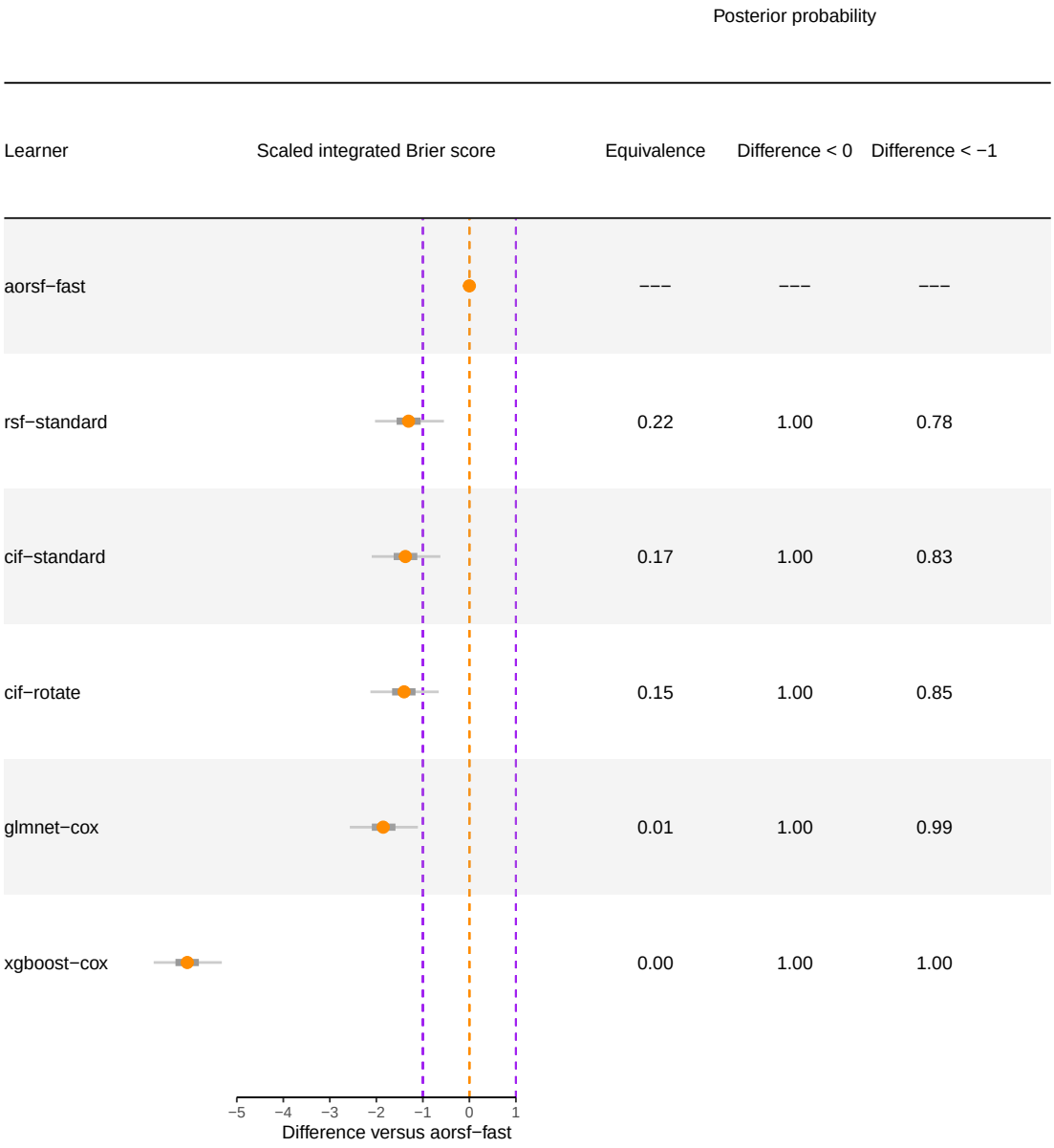
4.1.3 Data sets

We use a collection of publicly available data sets to benchmark the prediction accuracy and computational efficiency of the accelerated ORSF and each of the other learners described in Section 4.1.1

4.1.4 Statistical analysis

4.1.5 Results


xgboost−cox
glmnet−cox
cif−rotate
cif−standard
rsf−standard
aorsf−fast
Dataset; outcome
Index of Prediction Accuracy
Overall
ACTG 320; death
Monoclonal gammopathy; malignancy
FCL; relapse
ACTG 320; AIDS diagnosis
MESA; stroke
Heart Transplant; graft−loss or death
MESA; coronary heart disease
Early breast cancer; recurrence or death
GUIDE−IT; CVD death
SPRINT; CVD death
GUIDE−IT; HF hospitalization
ARIC; stroke
Colon cancer; recurrence
Colon cancer; death
FCL; death
MESA; heart failure
GBSG II; recurrence or death
SPRINT; death
Systolic Heart Failure; death
ARIC; coronary heart disease
Rotterdam tumor bank; recurrence
MESA; death
Rotterdam tumor bank; death
Monoclonal gammopathy; death
VA lung cancer trial; death
ARIC; death
ARIC; heart failure
Non−alcohol fatty liver disease; death
Serum free light chain; death
Primary biliary cholangitis; death
Movies released in 2015−2018; gross 1M USD
+1.2 (10%)
+1.1 (−1,361%)
+0.04 (1.7%)
+0.40 (7.9%)
+1.2 (19%)
+1.1 (16%)
+0.15 (1.7%)
+0.88 (9.7%)
+0.31 (3.1%)
+0.01 (0.09%)
+0.40 (2.6%)
+0.30 (2.1%)
+0.13 (0.84%)
+0.20 (1.0%)
+0.39 (1.7%)
+0.48 (2.3%)
+0.28 (1.1%)
+2.5 (6.2%)
0
10
20
30
40
50
60



4.2 Benchmark of variable importance

The aim of this numeric experiment is to compare negation VI to existing methods for VI. We consider methods that are intrinsic to the oblique RF (for exampleANOVA VI), those

that are intrinsic to the RF (for examplepermutation VI), and those that are model-agnostic (for exampleSHAP VI). All methods are evaluated based on their ability to discriminate between variables that are related to a censored outcome versus those that are not related to it.

4.2.1 Variable importance techniques

We compute permutation VI using the `randomForestSRC` package. Although the `party` package implements the approach to VI developed by Strobl et al. (2007), the developers of the `party` package note that the implementation of this approach for survival outcomes is "extremely slow and experimental" as of version 1.3.10. Therefore, it is not incorporated in the current study's numeric experiments. We compute ANOVA VI using the `aorsf` package, applying a p-value threshold of 0.01 as recommended by Menze et al. (2011). We compute SHAP VI for boosted tree models using the `xgboost` package and also compute SHAP VI for accelerated oblique RSFs using the `fastshap` package.

4.2.2 Variable types

We considered five classes of predictor variable, with each class characterized by its variables' relationship to a right-censored outcome. Specifically,

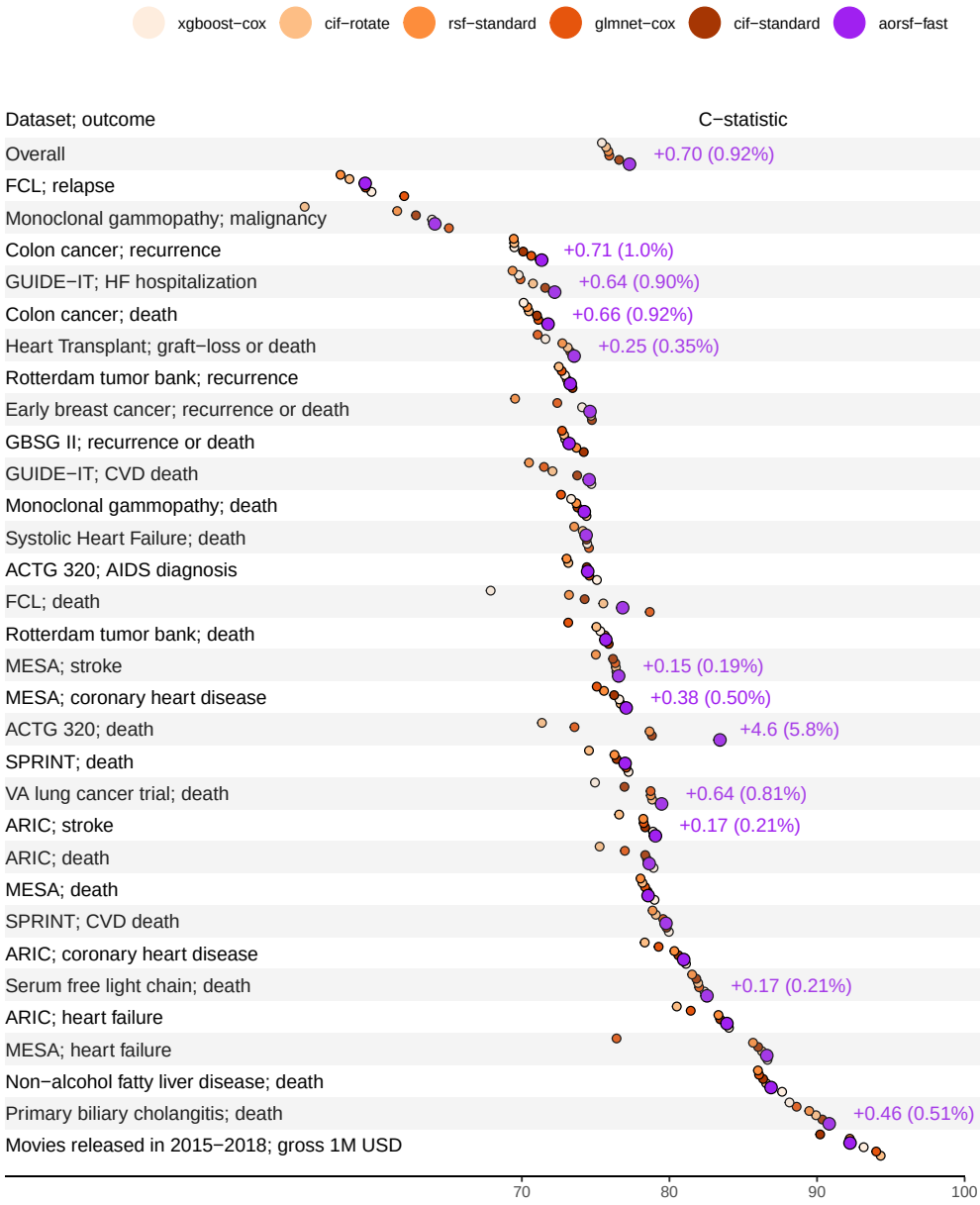
- *junk* variables had no relationship with the outcome.
- *main effect* variables were linearly related to the outcome
- *non-linear effect* variables had a sinusoidal relationship to the outcome
- *combination effect* variables had no direct relationship to the outcome. However, latent variables formed by combining the observed combination effects have a linear relationship to the outcome.
- *interaction effect* variables related to the outcome by multiplicative interaction with one other variable.

the variables have (1) no effect, (2) linear effect, (3) non-linear effect, (4) conditional linear effect, (5) combination linear effect, or (6) categorical effect on the outcome. Each conditional linear variable has an effect on the outcome that is modified by the value of one linear effect variable. Combination linear effect variables contribute to latent variables that are linearly related to the outcome. Categorical effect variables have 3 unique values, each with a different expected risk for the event of interest.

Acknowledgments

Research reported in this publication was supported by the Center for Biomedical Informatics, Wake Forest University School of Medicine. The project described was supported by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR001420. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Appendix A.


xgboost−cox
cif−rotate
rsf−standard
glmnet−cox
cif−standard
aorsf−fast
Dataset; outcome
C−statistic
Overall
FCL; relapse
Monoclonal gammopathy; malignancy
Colon cancer; recurrence
GUIDE−IT; HF hospitalization
Colon cancer; death
Heart Transplant; graft−loss or death
Rotterdam tumor bank; recurrence
Early breast cancer; recurrence or death
GBSG II; recurrence or death
GUIDE−IT; CVD death
Monoclonal gammopathy; death
Systolic Heart Failure; death
ACTG 320; AIDS diagnosis
FCL; death
Rotterdam tumor bank; death
MESA; stroke
MESA; coronary heart disease
ACTG 320; death
SPRINT; death
VA lung cancer trial; death
ARIC; stroke
ARIC; death
MESA; death
SPRINT; CVD death
ARIC; coronary heart disease
Serum free light chain; death
ARIC; heart failure
MESA; heart failure
Non−alcohol fatty liver disease; death
Primary biliary cholangitis; death
Movies released in 2015−2018; gross 1M USD
+0.70 (0.92%)
+0.71 (1.0%)
+0.64 (0.90%)
+0.66 (0.92%)
+0.25 (0.35%)
+0.15 (0.19%)
+0.38 (0.50%)
+4.6 (5.8%)
+0.64 (0.81%)
+0.17 (0.21%)
+0.17 (0.21%)
+0.46 (0.51%)
70
80
90
100

	Performance metric (SD)		Computing time, seconds			
			Fit model		Predict risk	
	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
Overall						
aorsf-fast	0.132 (0.114)	0.773 (0.071)	0.776	1.00	0.282	1.00
rsf-standard	0.120 (0.117)	0.759 (0.075)	2.016	2.60	0.444	1.57
cif-standard	0.119 (0.101)	0.766 (0.069)	4.663	6.01	18.044	63.9
cif-rotate	0.119 (0.130)	0.757 (0.080)	40.258	51.9	27.552	97.6
glmnet-cox	0.114 (0.124)	0.759 (0.076)	0.737	0.949	0.008	0.028
xgboost-cox	0.072 (0.103)	0.754 (0.089)	4.296	5.54	0.009	0.032
ACTG 320; AIDS diagnosis, n = 1151, p = 12						
cif-standard	0.024 (0.031)	0.744 (0.040)	1.580	10.7	3.948	116.0
aorsf-fast	0.024 (0.028)	0.745 (0.044)	0.147	1.00	0.034	1.00
glmnet-cox	0.016 (0.030)	0.746 (0.037)	0.161	1.10	0.002	0.059
rsf-standard	0.005 (0.041)	0.730 (0.042)	0.142	0.963	0.063	1.84
cif-rotate	0.004 (0.040)	0.731 (0.038)	15.182	103.2	4.149	121.9
xgboost-cox	0.000 (0.044)	0.751 (0.033)	3.630	24.7	0.003	0.088
ACTG 320; death, n = 1151, p = 12						
aorsf-fast	0.010 (0.020)	0.834 (0.057)	0.088	1.00	0.020	1.00
cif-standard	-0.001 (0.025)	0.788 (0.061)	1.589	18.1	3.861	192.9
xgboost-cox	-0.003 (0.004)	0.500 (0.000)	0.112	1.28	0.002	0.100
rsf-standard	-0.022 (0.053)	0.787 (0.071)	0.084	0.960	0.035	1.75
cif-rotate	-0.031 (0.048)	0.714 (0.084)	13.331	152.2	3.293	164.5
glmnet-cox	-0.053 (0.092)	0.736 (0.110)	0.262	2.99	0.002	0.100
ARIC; coronary heart disease, n = 13626, p = 31						
aorsf-fast	0.160 (0.008)	0.810 (0.007)	4.995	1.00	16.402	1.00
rsf-standard	0.156 (0.008)	0.803 (0.007)	11.043	2.21	1.245	0.076
cif-standard	0.134 (0.006)	0.806 (0.007)	68.008	13.6	441.507	26.9
glmnet-cox	0.124 (0.016)	0.793 (0.009)	1.354	0.271	0.822	0.050
cif-rotate	0.105 (0.003)	0.783 (0.008)	412.625	82.6	251.226	15.3
xgboost-cox	0.058 (0.016)	0.811 (0.007)	10.324	2.07	1.073	0.065
ARIC; death, n = 13626, p = 31						
rsf-standard	0.214 (0.007)	0.785 (0.004)	16.361	1.72	1.636	0.018
aorsf-fast	0.211 (0.008)	0.786 (0.005)	9.521	1.00	92.722	1.00
cif-standard	0.201 (0.005)	0.784 (0.005)	66.585	6.99	585.928	6.32
glmnet-cox	0.181 (0.019)	0.770 (0.008)	2.555	0.268	1.142	0.012
cif-rotate	0.150 (0.008)	0.753 (0.007)	427.333	44.9	524.889	5.66

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
xgboost-cox	0.123 (0.006)	0.789 (0.004)	10.891	1.14	1.028	0.011
ARIC; heart failure, n = 13626, p = 31						
aorsf-fast	0.236 (0.008)	0.839 (0.006)	5.758	1.00	21.727	1.00
rsf-standard	0.232 (0.009)	0.833 (0.006)	11.647	2.02	1.259	0.058
cif-standard	0.202 (0.006)	0.834 (0.006)	71.819	12.5	476.484	21.9
cif-rotate	0.171 (0.005)	0.805 (0.008)	431.960	75.0	319.739	14.7
glmnet-cox	0.163 (0.041)	0.814 (0.016)	2.435	0.423	1.106	0.051
xgboost-cox	0.117 (0.019)	0.840 (0.005)	11.828	2.05	1.131	0.052
ARIC; stroke, n = 13626, p = 31						
aorsf-fast	0.091 (0.006)	0.791 (0.010)	5.070	1.00	12.402	1.00
rsf-standard	0.090 (0.007)	0.782 (0.011)	10.586	2.09	1.225	0.099
glmnet-cox	0.076 (0.006)	0.783 (0.010)	1.135	0.224	0.905	0.073
cif-standard	0.073 (0.005)	0.784 (0.011)	69.781	13.8	389.705	31.4
cif-rotate	0.052 (0.003)	0.766 (0.010)	407.352	80.4	170.815	13.8
xgboost-cox	0.048 (0.023)	0.789 (0.010)	7.716	1.52	0.915	0.074
Colon cancer; death, n = 929, p = 12						
aorsf-fast	0.100 (0.014)	0.718 (0.011)	0.259	1.00	0.075	1.00
cif-standard	0.097 (0.013)	0.710 (0.012)	1.193	4.60	3.848	51.3
cif-rotate	0.087 (0.017)	0.705 (0.014)	13.358	51.5	5.116	68.1
rsf-standard	0.087 (0.019)	0.704 (0.011)	1.226	4.73	0.150	2.00
glmnet-cox	0.076 (0.016)	0.711 (0.019)	0.116	0.448	0.004	0.053
xgboost-cox	0.063 (0.013)	0.701 (0.013)	3.165	12.2	0.004	0.053
Colon cancer; recurrence, n = 929, p = 12						
aorsf-fast	0.100 (0.017)	0.713 (0.016)	0.259	1.00	0.076	1.00
cif-standard	0.091 (0.015)	0.701 (0.017)	1.182	4.56	3.929	51.7
cif-rotate	0.084 (0.020)	0.695 (0.017)	13.338	51.5	5.118	67.3
rsf-standard	0.081 (0.020)	0.694 (0.015)	1.223	4.72	0.151	1.99
glmnet-cox	0.073 (0.018)	0.706 (0.024)	0.115	0.444	0.004	0.053
xgboost-cox	0.060 (0.011)	0.695 (0.018)	2.895	11.2	0.004	0.053
Early breast cancer; recurrence or death, n = 614, p = 1692						
cif-rotate	0.070 (0.018)	0.747 (0.027)	6304.881	8,318.3	374.594	2,134.8
cif-standard	0.067 (0.019)	0.747 (0.030)	8.059	10.6	4.493	25.6
aorsf-fast	0.065 (0.028)	0.746 (0.026)	0.758	1.00	0.175	1.00
glmnet-cox	0.041 (0.032)	0.724 (0.036)	5.686	7.50	0.005	0.029
xgboost-cox	0.027 (0.034)	0.741 (0.030)	2.154	2.84	0.006	0.034
rsf-standard	0.024 (0.037)	0.695 (0.033)	0.323	0.427	0.666	3.80

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
FCL; death, n = 541, p = 7						
glmnet-cox	0.117 (0.028)	0.787 (0.037)	0.082	0.988	0.002	0.111
aorsf-fast	0.100 (0.038)	0.768 (0.033)	0.083	1.00	0.018	1.00
cif-rotate	0.087 (0.048)	0.755 (0.027)	6.369	76.7	1.766	98.0
cif-standard	0.084 (0.038)	0.743 (0.036)	0.751	9.04	1.020	56.6
rsf-standard	0.072 (0.048)	0.732 (0.034)	0.589	7.08	0.040	2.22
xgboost-cox	0.029 (0.050)	0.679 (0.121)	0.320	3.86	0.002	0.111
FCL; relapse, n = 541, p = 7						
glmnet-cox	0.029 (0.017)	0.620 (0.024)	0.086	0.717	0.003	0.120
xgboost-cox	0.010 (0.016)	0.598 (0.032)	1.272	10.6	0.003	0.120
cif-standard	0.007 (0.021)	0.594 (0.023)	0.753	6.27	1.153	46.1
aorsf-fast	0.007 (0.019)	0.594 (0.025)	0.120	1.00	0.025	1.00
cif-rotate	-0.012 (0.025)	0.583 (0.030)	7.360	61.3	2.955	118.1
rsf-standard	-0.027 (0.032)	0.577 (0.024)	0.742	6.18	0.092	3.68
GBSG II; recurrence or death, n = 686, p = 10						
cif-standard	0.123 (0.020)	0.742 (0.019)	0.919	5.21	2.297	52.1
rsf-standard	0.120 (0.023)	0.737 (0.019)	1.198	6.79	0.115	2.61
aorsf-fast	0.116 (0.024)	0.732 (0.017)	0.176	1.00	0.044	1.00
cif-rotate	0.107 (0.022)	0.729 (0.017)	11.084	62.8	3.797	86.2
glmnet-cox	0.091 (0.019)	0.727 (0.021)	0.094	0.533	0.003	0.068
xgboost-cox	0.083 (0.017)	0.729 (0.020)	2.257	12.8	0.003	0.068
GUIDE-IT; CVD death, n = 894, p = 59						
aorsf-fast	0.075 (0.018)	0.746 (0.027)	0.166	1.00	0.036	1.00
glmnet-cox	0.063 (0.041)	0.715 (0.091)	0.995	5.99	0.002	0.056
cif-rotate	0.059 (0.016)	0.721 (0.025)	35.119	211.4	4.542	126.1
cif-standard	0.058 (0.014)	0.737 (0.022)	1.771	10.7	3.109	86.3
rsf-standard	0.046 (0.023)	0.705 (0.025)	0.638	3.84	0.063	1.75
xgboost-cox	0.039 (0.051)	0.747 (0.020)	4.202	25.3	0.003	0.083
GUIDE-IT; HF hospitalization, n = 894, p = 59						
aorsf-fast	0.082 (0.019)	0.722 (0.025)	0.262	1.00	0.551	1.00
cif-standard	0.071 (0.010)	0.716 (0.023)	1.796	6.85	3.342	6.07
cif-rotate	0.068 (0.018)	0.708 (0.029)	41.195	157.1	5.392	9.79
rsf-standard	0.059 (0.022)	0.694 (0.026)	1.230	4.69	0.123	0.224
glmnet-cox	0.058 (0.019)	0.699 (0.024)	0.914	3.49	0.003	0.005
xgboost-cox	0.039 (0.017)	0.698 (0.027)	2.792	10.6	0.003	0.005
Heart Transplant; graft-loss or death, n = 3787, p = 52						

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
cif-rotate	0.056 (0.010)	0.731 (0.017)	152.873	160.7	31.205	107.1
aorsf-fast	0.052 (0.006)	0.735 (0.014)	0.951	1.00	0.291	1.00
cif-standard	0.050 (0.006)	0.733 (0.013)	10.380	10.9	38.621	132.6
rsf-standard	0.048 (0.009)	0.727 (0.013)	2.246	2.36	0.893	3.07
glmnet-cox	0.036 (0.006)	0.711 (0.016)	0.920	0.967	0.018	0.062
xgboost-cox	0.031 (0.007)	0.716 (0.018)	3.383	3.56	0.011	0.038
MESA; coronary heart disease, n = 6785, p = 28						
aorsf-fast	0.055 (0.009)	0.771 (0.014)	1.720	1.00	0.520	1.00
rsf-standard	0.051 (0.011)	0.756 (0.016)	3.294	1.92	0.430	0.828
cif-standard	0.049 (0.006)	0.763 (0.014)	20.168	11.7	102.381	197.0
cif-rotate	0.047 (0.006)	0.767 (0.017)	162.455	94.4	39.029	75.1
glmnet-cox	0.031 (0.017)	0.751 (0.025)	0.751	0.436	0.014	0.027
xgboost-cox	0.025 (0.027)	0.766 (0.016)	5.609	3.26	0.018	0.035
MESA; death, n = 6793, p = 28						
aorsf-fast	0.144 (0.009)	0.786 (0.010)	1.797	1.00	1.069	1.00
rsf-standard	0.141 (0.009)	0.780 (0.010)	3.613	2.01	0.487	0.456
cif-standard	0.136 (0.007)	0.785 (0.010)	20.306	11.3	114.250	106.9
cif-rotate	0.126 (0.007)	0.782 (0.010)	171.028	95.2	74.574	69.8
glmnet-cox	0.123 (0.025)	0.783 (0.013)	0.702	0.391	0.046	0.043
xgboost-cox	0.054 (0.027)	0.790 (0.010)	10.187	5.67	0.052	0.049
MESA; heart failure, n = 6785, p = 28						
aorsf-fast	0.108 (0.010)	0.866 (0.014)	1.619	1.00	0.427	1.00
rsf-standard	0.108 (0.013)	0.857 (0.015)	3.844	2.37	0.479	1.12
cif-standard	0.099 (0.008)	0.860 (0.016)	20.168	12.5	100.439	235.4
cif-rotate	0.098 (0.009)	0.863 (0.014)	148.543	91.7	36.415	85.4
glmnet-cox	0.036 (0.037)	0.764 (0.137)	0.893	0.551	0.019	0.045
xgboost-cox	-0.002 (0.025)	0.866 (0.012)	8.241	5.09	0.016	0.038
MESA; stroke, n = 6783, p = 28						
aorsf-fast	0.025 (0.006)	0.766 (0.015)	1.689	1.00	0.393	1.00
cif-rotate	0.025 (0.005)	0.764 (0.016)	151.593	89.7	32.758	83.4
cif-standard	0.025 (0.004)	0.762 (0.015)	19.936	11.8	98.802	251.4
rsf-standard	0.022 (0.008)	0.750 (0.013)	2.453	1.45	0.737	1.87
glmnet-cox	0.022 (0.008)	0.763 (0.016)	0.775	0.459	0.012	0.031
xgboost-cox	0.008 (0.026)	0.764 (0.015)	4.517	2.67	0.015	0.038
Monoclonal gammopathy; death, n = 1384, p = 8						
cif-rotate	0.160 (0.019)	0.744 (0.014)	15.319	34.5	6.540	62.8
aorsf-fast	0.158 (0.016)	0.742 (0.011)	0.444	1.00	0.104	1.00

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
cif-standard	0.151 (0.015)	0.738 (0.012)	1.432	3.23	6.299	60.5
rsf-standard	0.151 (0.017)	0.737 (0.011)	1.906	4.29	0.207	1.99
glmnet-cox	0.138 (0.021)	0.727 (0.014)	0.131	0.295	0.004	0.038
xgboost-cox	0.123 (0.012)	0.733 (0.012)	3.450	7.77	0.004	0.038
Monoclonal gammopathy; malignancy, n = 1384, p = 8						
glmnet-cox	0.015 (0.011)	0.651 (0.055)	0.118	0.599	0.002	0.050
aorsf-fast	0.010 (0.014)	0.641 (0.036)	0.197	1.00	0.040	1.00
xgboost-cox	0.007 (0.017)	0.639 (0.039)	1.734	8.79	0.003	0.075
cif-standard	0.006 (0.011)	0.628 (0.033)	1.443	7.32	5.801	144.9
rsf-standard	-0.009 (0.018)	0.616 (0.036)	0.684	3.47	0.077	1.92
cif-rotate	-0.024 (0.023)	0.553 (0.035)	12.997	65.9	4.817	120.3
Movies released in 2015-2018; gross 1M USD, n = 551, p = 46						
cif-rotate	0.639 (0.023)	0.943 (0.007)	20.129	26.2	5.703	105.5
glmnet-cox	0.620 (0.033)	0.940 (0.009)	0.671	0.872	0.003	0.056
rsf-standard	0.520 (0.021)	0.922 (0.010)	1.666	2.17	0.103	1.91
aorsf-fast	0.517 (0.026)	0.922 (0.012)	0.769	1.00	0.054	1.00
xgboost-cox	0.511 (0.028)	0.932 (0.009)	16.829	21.9	0.005	0.093
cif-standard	0.475 (0.028)	0.902 (0.018)	0.867	1.13	2.080	38.5
Non-alcohol fatty liver disease; death, n = 17549, p = 24						
aorsf-fast	0.215 (0.009)	0.869 (0.005)	5.813	1.00	11.773	1.00
rsf-standard	0.210 (0.009)	0.860 (0.005)	10.989	1.89	1.393	0.118
glmnet-cox	0.209 (0.011)	0.861 (0.005)	2.110	0.363	0.124	0.011
cif-standard	0.208 (0.007)	0.863 (0.006)	64.081	11.0	674.392	57.3
cif-rotate	0.193 (0.008)	0.866 (0.005)	254.040	43.7	194.487	16.5
xgboost-cox	0.023 (0.015)	0.876 (0.005)	8.628	1.48	1.012	0.086
Primary biliary cholangitis; death, n = 276, p = 19						
aorsf-fast	0.430 (0.033)	0.908 (0.021)	0.075	1.00	0.018	1.00
cif-rotate	0.405 (0.041)	0.899 (0.022)	9.498	126.5	2.287	126.9
rsf-standard	0.391 (0.034)	0.895 (0.023)	0.102	1.36	0.039	2.17
cif-standard	0.351 (0.034)	0.904 (0.025)	0.194	2.59	0.850	47.1
glmnet-cox	0.342 (0.045)	0.886 (0.028)	0.104	1.39	0.002	0.111
xgboost-cox	0.256 (0.104)	0.881 (0.027)	4.587	61.1	0.003	0.167
Rotterdam tumor bank; death, n = 2982, p = 11						
aorsf-fast	0.162 (0.012)	0.757 (0.009)	1.488	1.00	0.397	1.00
cif-standard	0.160 (0.010)	0.759 (0.009)	4.838	3.25	29.075	73.2
rsf-standard	0.160 (0.014)	0.756 (0.009)	2.974	2.00	0.846	2.13
cif-rotate	0.148 (0.011)	0.750 (0.011)	34.842	23.4	32.884	82.8

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
xgboost-cox	0.131 (0.014)	0.753 (0.010)	4.263	2.87	0.017	0.043
glmnet-cox	0.119 (0.008)	0.731 (0.009)	0.779	0.523	0.019	0.048
Rotterdam tumor bank; recurrence, n = 2982, p = 11						
cif-standard	0.145 (0.011)	0.734 (0.009)	4.883	2.98	29.781	67.9
aorsf-fast	0.143 (0.011)	0.733 (0.009)	1.636	1.00	0.438	1.00
rsf-standard	0.139 (0.012)	0.731 (0.008)	3.160	1.93	0.836	1.91
cif-rotate	0.130 (0.010)	0.725 (0.009)	36.274	22.2	36.461	83.2
glmnet-cox	0.118 (0.008)	0.727 (0.008)	0.815	0.498	0.029	0.066
xgboost-cox	0.114 (0.008)	0.729 (0.009)	3.774	2.31	0.020	0.046
Serum free light chain; death, n = 7874, p = 10						
aorsf-fast	0.252 (0.014)	0.825 (0.007)	3.093	1.00	4.842	1.00
glmnet-cox	0.249 (0.012)	0.820 (0.007)	1.261	0.408	0.081	0.017
cif-standard	0.245 (0.011)	0.818 (0.008)	19.554	6.32	146.176	30.2
rsf-standard	0.245 (0.013)	0.815 (0.008)	6.058	1.96	0.611	0.126
cif-rotate	0.230 (0.009)	0.819 (0.007)	67.331	21.8	132.391	27.3
xgboost-cox	0.097 (0.038)	0.824 (0.007)	6.139	1.98	0.084	0.017
SPRINT; CVD death, n = 9361, p = 174						
glmnet-cox	0.072 (0.011)	0.796 (0.011)	12.974	4.88	0.026	0.025
aorsf-fast	0.070 (0.006)	0.798 (0.011)	2.657	1.00	1.054	1.00
rsf-standard	0.065 (0.007)	0.789 (0.014)	3.854	1.45	0.701	0.665
cif-standard	0.062 (0.003)	0.798 (0.011)	48.137	18.1	185.155	175.7
cif-rotate	0.061 (0.005)	0.791 (0.012)	907.261	341.4	135.081	128.2
xgboost-cox	0.006 (0.017)	0.800 (0.011)	6.905	2.60	0.029	0.028
SPRINT; death, n = 9361, p = 174						
glmnet-cox	0.124 (0.012)	0.771 (0.009)	4.886	1.02	0.065	0.017
aorsf-fast	0.117 (0.008)	0.770 (0.008)	4.807	1.00	3.794	1.00
rsf-standard	0.111 (0.008)	0.763 (0.009)	6.757	1.41	0.752	0.198
cif-standard	0.107 (0.006)	0.764 (0.008)	48.766	10.1	208.039	54.8
cif-rotate	0.092 (0.007)	0.745 (0.009)	1024.991	213.2	194.183	51.2
xgboost-cox	0.031 (0.023)	0.772 (0.008)	9.163	1.91	0.077	0.020
Systolic Heart Failure; death, n = 2231, p = 41						
glmnet-cox	0.114 (0.013)	0.746 (0.012)	0.257	0.387	0.009	0.040
cif-rotate	0.113 (0.013)	0.741 (0.011)	69.125	104.0	18.805	84.2
aorsf-fast	0.110 (0.015)	0.744 (0.011)	0.665	1.00	0.223	1.00
cif-standard	0.110 (0.011)	0.744 (0.011)	3.672	5.53	18.416	82.5
rsf-standard	0.105 (0.011)	0.735 (0.011)	2.072	3.12	0.298	1.34
xgboost-cox	0.090 (0.009)	0.744 (0.010)	4.047	6.09	0.009	0.040

(continued)

	Scaled Brier	C-Statistic	Median	Ratio	Median	Ratio
VA lung cancer trial; death, n = 137, p = 8						
aorsf-fast	0.200 (0.050)	0.795 (0.034)	0.052	1.00	0.011	1.00
cif-rotate	0.198 (0.066)	0.788 (0.036)	4.620	88.8	1.017	92.3
rsf-standard	0.175 (0.048)	0.787 (0.037)	0.063	1.21	0.035	3.18
glmnet-cox	0.160 (0.036)	0.787 (0.038)	0.075	1.44	0.002	0.182
cif-standard	0.128 (0.040)	0.770 (0.037)	0.093	1.79	0.133	12.1
xgboost-cox	0.068 (0.078)	0.750 (0.046)	1.144	22.0	0.002	0.182

References

Paul Blanche, Jean-François Dartigues, and Hélène Jacqmin-Gadda. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in medicine*, 32(30):5381–5397, 2013.

Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

Erika Graf, Claudia Schmoor, Willi Sauerbrei, and Martin Schumacher. Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine*, 18(17-18):2529–2545, 1999. URL `https://doi.org/10.1002/(SICI)1097-0258(19990915/30)18:17/18%3C2529::AID-SIM274%3E3.0.CO;2-5`.

Frank E. Harrell, Robert M. Califf, David B. Pryor, Kerry L. Lee, and Robert A. Rosati. Evaluating the Yield of Medical Tests. *JAMA*, 247(18):2543–2546, 05 1982. ISSN 0098-7484. doi: 10.1001/jama.1982.03320430047030. URL `https://doi.org/10.1001/jama.1982.03320430047030`.

Torsten Hothorn, Berthold Lausen, Axel Benner, and Martin Radespiel-Tröger. Bagging survival trees. *Statistics in medicine*, 23(1):77–91, 2004.

Torsten Hothorn, Kurt Hornik, and Achim Zeileis. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics*, 15 (3):651–674, 2006.

Torsten Hothorn, Kurt Hornik, Carolin Strobl, and Achim Zeileis. Party: a laboratory for recursive partytioning, 2010.

H. Ishwaran and U.B. Kogalur. *Random Forests for Survival, Regression, and Classification (RF-SRC)*, 2019. URL `https://cran.r-project.org/package=randomForestSRC`. R package version 2.8.0, available at `https://cran.r-project.org/package=randomForestSRC`.

Hemant Ishwaran and Udaya B Kogalur. Consistency of random survival forests. *Statistics & probability letters*, 80(13-14):1056–1064, 2010.

Hemant Ishwaran, Udaya B Kogalur, Eugene H Blackstone, and Michael S Lauer. Random survival forests. *The Annals of Applied Statistics*, pages 841–860, 2008.

Byron C Jaeger, D Leann Long, Dustin M Long, Mario Sims, Jeff M Szychowski, Yuan-I Min, Leslie A Mcclure, George Howard, and Noah Simon. Oblique random survival forests. *The Annals of Applied Statistics*, 13(3):1847–1883, 2019.

Michael W Kattan and Thomas A Gerds. The index of prediction accuracy: an intuitive measure useful for evaluating risk prediction models. *Diagnostic and prognostic research*, 2(1):1–7, 2018.

Rakesh Katuwal, Ponnuthurai Nagaratnam Suganthan, and Le Zhang. Heterogeneous oblique random forest. *Pattern Recognition*, 99:107078, 2020.

Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

Bjoern H Menze, B Michael Kelm, Daniel N Splitthoff, Ullrich Koethe, and Fred A Hamprecht. On oblique random forests. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 453–469. Springer, 2011.

Karel GM Moons, Andre Pascal Kengne, Diederick E Grobbee, Patrick Royston, Yvonne Vergouwe, Douglas G Altman, and Mark Woodward. Risk prediction models: II. external validation, model updating, and impact assessment. *Heart*, 98(9):691–698, 2012a.

Karel GM Moons, Andre Pascal Kengne, Mark Woodward, Patrick Royston, Yvonne Vergouwe, Douglas G Altman, and Diederick E Grobbee. Risk prediction models: I. development, internal validation, and assessing the incremental value of a new (bio) marker. *Heart*, 98(9):683–690, 2012b.

Nitesh Poona, Adriaan Van Niekerk, and Riyad Ismail. Investigating the utility of oblique tree-based ensembles for the classification of hyperspectral data. *Sensors*, 16(11):1918, 2016.

Xueheng Qiu, Le Zhang, Ponnuthurai Nagaratnam Suganthan, and Gehan AJ Amaratunga. Oblique random forest ensemble via least square estimation for time series forecasting. *Information Sciences*, 420:249–262, 2017.

Tom Rainforth and Frank Wood. Canonical correlation forests. *arXiv preprint arXiv:1507.05444*, 2015.

Carolin Strobl, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics*, 8(1):25, 2007.

Terry Therneau. Survival package source code documentation, April 2022. URL https://github.com/therneau/survival/blob/5440691d44abea537b08aeb60153a31654d66a9b/noweb. original-date: 2016-04-28.

Tyler M Tomita, James Browne, Cencheng Shen, Jaewon Chung, Jesse L Patsolic, Benjamin Falk, Carey E Priebe, Jason Yim, Randal Burns, Mauro Maggioni, et al. Sparse projection oblique randomer forests. *Journal of machine learning research*, 21(104), 2020.

Hong Wang and Gang Li. A selective review on random survival forests for high dimensional data. *Quantitative bio-science*, 36(2):85, 2017.

Hong Wang and Lifeng Zhou. Random survival forest with space extensions for censored data. *Artificial intelligence in medicine*, 79:52–61, 2017.

Marvin N. Wright and Andreas Ziegler. ranger: A fast implementation of random forests for high dimensional data in c++ and r. *Journal of Statistical Software*, 77(1):1–17, 2017. doi: 10.18637/jss.v077.i01. URL https://www.jstatsoft.org/index.php/jss/article/view/v077i01.

Le Zhang and Ponnuthurai N Suganthan. Oblique decision tree ensemble via multisurface proximal support vector machine. *IEEE transactions on cybernetics*, 45(10):2165–2176, 2014.

Lifeng Zhou, Hong Wang, and Qingsong Xu. Random rotation survival forest for high dimensional censored data. *SpringerPlus*, 5(1):1–10, 2016.

Ruoqing Zhu. *Tree-based Methods for Survival Analysis and High-dimensional Data*. PhD thesis, The University of North Carolina at Chapel Hill, 2013.

Ruoqing Zhu, Donglin Zeng, and Michael R Kosorok. Reinforcement learning trees. *Journal of the American Statistical Association*, 110(512):1770–1784, 2015.